UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Identification of patterns in control charts for processes with statistically correlated noise

De La Torre Gutierrez, Hector; Pham, Duc

DOI:
10.1080/00207543.2017.1360530

License:
None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
De La Torre Gutierrez, H & Pham, D 2017, 'Identification of patterns in control charts for processes with statistically correlated noise', *International Journal of Production Research*.
https://doi.org/10.1080/00207543.2017.1360530

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 05. May. 2023

Identification of patterns in control charts for processes with statistically correlated noise

Authors (FAMILY NAME, GIVEN NAME):

De la Torre Gutiérrez, Héctor*[1]

Pham, Duc Truong[2]

AFFILIATION:

[1,2] School of Engineering, University of Birmingham,
Edgbaston, Birmingham, West Midlands B15 2TT

*Corresponding author (Email: hxd394@bham.ac.uk)

For Peer Review Only

Abstract

In real industrial scenarios, if the quality characteristics of a continuous or batch production process are monitored using Shewhart control charts, there could be a large number of false alarms about the process going out of control. This is because these control charts assume that the inherent noise of the monitored process is normally, independently and identically distributed, although the assumption of independence is not always correct for continuous and batch production processes. This paper presents three control chart pattern recognition systems where the inherent disturbance is assumed to be stationary. The systems use the first-order autoregressive (AR(1)), moving-average (MA(1)) and autoregressive moving-average (ARMA(1,1)) models. A special pattern generation scheme is adopted to ensure generality, randomness, and comparability, as well as allowing the further categorisation of the studied patterns. Two different input representation techniques for the recognition systems were studied. These gave nearly the same performance for the MA(1) and ARMA(1,1) models, while the raw data yielded the highest accuracies when AR(1) was used. The effect of autocorrelation on the pattern recognition capabilities of the developed models was studied. It was observed that Normal and Upward Shift patterns were the most affected.

Keywords: Control chart pattern recognition, stationary processes, dynamic regression, pattern generation, Support Vector Machine

For Peer Review Only

Nomenclature

AR	Autoregressive
ARMA	Autoregressive moving-average
BPNN	Back-propagation neural network
BESS	Bessel kernel
CCPR	Control chart pattern recognition
CYC	Cyclic pattern
D_t	Disturbance magnitude at time *t*
DS	Downward shift pattern
DT	Downward trend pattern
e_t	Random error (white-noise) at time *t*
IMA	Integrated moving-average model
IRT	Input representation technique
LAPLA	Laplace kernel
LVQ	Learning Vector quantisation
MA	Moving-average model
ML	Machine learning
N_t	Inherent disturbance at time t
NIID	Normally, identically and independently distributed
NLM	Nonlinear model
NLM-ARMA	Nonlinear regression model with autoregressive moving-average errors
NORM	Normal pattern
PGS	Pattern generation scheme
r_1	Autocorrelation coefficient
RBF	Radial basis function kernel

For Peer Review Only

S_Y	Standard deviation of the variable Y
sin	Sine function
SSE	Sum of squared error
SVM	Support Vector Machine
SYS	Systematic pattern
TANH	Hyperbolic Tangent kernel
US	Upward shift pattern
UT	Upward trend pattern
Y_t	Measurement of the quality characteristic under study at time *t*
Z_t	Scaled variable at time *t*
α	Significance level
β	Abnormal pattern parameter
σ_N	Standard deviation of the variable N
μ	Mean value
τ	Time when a break point occurs
θ	Moving-average coefficient
ϕ	Autoregressive coefficient

For Peer Review Only

1 Introduction

Control charts are graphical tools that monitor and assess the performance of production processes, revealing abnormal (deterministic) disturbances when there is a shift in the process. When the process is operating normally, a "Normal" control chart pattern can be observed (see Western Electric Company 1958). If an assignable cause is affecting the process, the control chart can exhibit one or more of fourteen types of patterns (Western Electric Company 1958), six of these being considered as basic patterns: Upward/Downward Trends (UT/DT), Upward/Downward Shifts (US/DS), Cyclic and Systematic (CYC and SYS) (see Western Electric Company 1958). The remaining eight patterns are either particular cases or combinations of these basic patterns.

The efficient and accurate identification of these simple patterns has been a problem studied by researchers during the last two decades. These researchers have focused mainly on the identification of patterns when the inherent disturbance of the process is purely random, i.e. normally, independently and identically distributed (NIID) (Kazemi et al. 2015; Chompu-Inwai & Thaiupathump 2015; Xanthopoulos & Razzaghi 2014). Unfortunately, the assumption of uncorrelated (independent) observations is not even approximately satisfied in some processes. Examples include chemical processes where consecutive measurements of process or product characteristics are often highly correlated, or automated test and inspection procedures, where every quality characteristic is measured on every unit in time order of production (Montgomery 2009).

Disturbances in control charts can be divided into two categories: stochastic and deterministic (MacGregor 1988). Stochastic disturbances are those related to fluctuation in the performance of the process with time, and in most cases, can be modelled by time series. In this paper, stochastic disturbances are assumed to be inherent to the process. On the other hand, deterministic disturbances are those that are independent of the process and are fully

For Peer Review Only

determined by external variables; these can be caused by faults in machines, tools or materials or by environmental variables or other factors. They manifest themselves as abnormal patterns in the control chart (Western Electric Company 1958).

Very few studies have been performed regarding control chart pattern recognition (CCPR) when the inherent disturbance of the process is autocorrelated, the first-order autoregressive model (AR(1)) being the only utilised so far. The aim of this work is to develop new CCPR systems where the inherent disturbance is modelled by other stationary time series.

Three time series models for stationary processes are used to model the inherent disturbances: AR(1), MA(1) and ARMA (1,1). Stationarity implies that the generated time series varies in a stable manner about a fixed mean (Box et al. 2011). In this work, independence is assumed between the causes of the inherent disturbance and the causes of patterns in the control chart, thus retaining the same seven patterns observed in NIID processes. The autocorrelation level is expected to remain constant during the pattern recognition window and is also assumed to be unknown.

The CCPR systems developed comprise a Machine Learning (ML) algorithm named the Support Vector Machine (SVM), as the pattern recognition system, and a pattern generation scheme (PGS) for first-order stationary processes. PGSs synthesise patterns for CCPR systems taking into account the main aim of identifying patterns in control charts: automate the analysis and identification of patterns in control charts, allowing a full categorisation of the fault.

The training of ML algorithms requires large amounts of data which are not readily available. It is therefore necessary to synthesise patterns that are representative of those to be recognised during the operation of the CCPR system.

For Peer Review Only

The PGS employed in this work is related to that proposed by De La Torre Gutierrez and Pham (2016) where the inherent noise is NIID. As the assumption of independence is not fulfilled in autocorrelated[1] processes such as those that are continuous, the PGS must be capable of generating patterns of this type, taking into account the real variance of the process. Under the assumption of independence between stochastic and deterministic disturbances, the dynamic regression model named Nonlinear regression model with autoregressive moving-average errors (NLM-ARMA) is able to represent the data generated by the PGS instead of the nonlinear regression model proposed by De La Torre Gutierrez and Pham (2016). This PGS scheme fulfils the conditions needed for the design of CCPR systems previously mentioned. The p-value (α) of the NLM-ARMA model fitted to the control chart data determines whether or not to categorise it in its initial class, thus making this PGS robust to variations in the training pattern parameters.

In the proposed PGS, unbiased estimation of common-cause parameters (autocorrelation coefficients, standard deviation and mean value) and pattern parameters is ensured. If the autocorrelation and pattern parameters are not estimated simultaneously, an abnormal pattern can badly affect their estimation, i.e., they can be biased due to the mixing of the effects of the two types of noise. For instance, Figure 1(a) shows a Normal pattern generated from an AR(1) process with autocorrelation coefficient ϕ=0.5. Figure 1(b) shows the same pattern added to a trend of slope $0.05\sigma_y$. Using equation (1) for estimating ϕ from Figure 1(a) data, the estimated value of autocorrelation is ϕ=0.50.

$$r_1 = \frac{\sum_{t=1}^{n-1}(Y_t - \bar{Y})(Y_{t+1} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2} \quad (1)$$

where r_1 represents the estimated value of the autocorrelation coefficient of the AR(1) process.

[1] In this work, autocorrelation will refer to both autoregressive and moving-average models.

For Peer Review Only

On the other hand, in the control chart shown in Figure 1(b), the estimated value of ϕ is 0.81. Thus, the estimation of ϕ is biased by the positive trend.

Such biasing of parameters when assignable causes are present was studied by Woodall and Faltin (1993) and Boyles (2000) who found that if pattern recognition is employed to detect assignable causes, estimators of common-cause, in this case the autocorrelation level, are highly biased when an abnormal pattern is present.

Figure 1: Two simple patterns with ϕ=0.50

In the literature, CCPR models that deal with the recognition and classification of significant patterns and that can estimate the corresponding parameters while minimising misclassification errors are very rare (Lesany et al. 2013). In this work, it is proposed to employ a PGS that is able to separate common-cause and assignable-cause effects and estimate them simultaneously. However, the estimation is carried out independently as the two disturbances are assumed to be independent due to their different sources. The NLM-ARMA model used in the proposed PGS ensures such conditions of parameter estimation (Pankratz 1991; Hyndman & Athanasopoulos 2014).

Once the synthetic patterns are generated by applying the proposed PGS, the next step is to train the ML algorithm. In this paper, three factors relating to the input of the CCPR system are studied: input representation technique (IRT), PGS and kernel of the pattern recognition system. Finding arrangements of these factors that achieve the highest accuracies when the inherent noise is autocorrelated is another aim of this research.

The remainder of the paper is organised as follows. Section 2 reviews the most relevant work related to this paper. The proposed CCPR is introduced in section 3. Section 4 presents the results obtained. Finally, conclusions and suggestions for future research are given in section 5.

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

2 Literature review

This section reviews work related to CCPR for autocorrelated processes, synthesis of patterns for training CCPR systems and estimation of parameters of control charts using CCPR systems.

2.1. Monitoring autocorrelated processes

The literature on monitoring autocorrelated processes is extensive. This subsection summarises work that is most relevant to this research.

There is a substantial discussion in the literature on choosing the disturbance model for the inherent noise of the process (Wang & Tsung 2007). Authors such as Zhang and Pollard (1994), Nembhard and Kao (2003) and Hwarng (2004) used the AR(1) model to describe the inherent disturbance of some industrial processes, Tsung et al. (1998), Jiang and Tsui (2002) and Wang and Tsung (2007) employed the ARMA (1,1) model. Montgomery et al. (2000) and Jiang and Tsui (2002) modelled the inherent disturbance using an integrated moving-average model (IMA (1,1)).

The most common procedure for monitoring autocorrelated processes consists of plotting the residuals of a fitted time series model. If the fitted model is adequate, these residuals will be NIID and thus traditional control charts can be used to monitor them. Recent research found that monitoring residuals affects the detection of mean shifts and highly depends on the ability to fit time series models and a-priori knowledge of the process (Longnecker and Ryan 1992; Zhang 1997; Lu and Reynolds 1999).

Recent advances have been made by forecast-based monitoring schemes to address this problem. Dyer et al. (2003) proposed a forecast-based monitoring scheme for three stationary processes, AR (1), MA (1) and ARMA(1,1). Alwan (1991) investigated the effect of

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

autocorrelation on masking the effect of special causes and also studied how run-rules and static control limits can increase the number of false alarms. Boyles (2000) studied the problem of splitting common-cause signals from assignable-cause signals by means of a standard estimator for first-order autocorrelated processes.

For a literature review and broader discussion of Statistical Process Control methods for monitoring autocorrelated processes, see Psarakis and Papaleonida (2007).

2.2. CCPR for NIID processes

The design of CCPR systems for processes where the inherent disturbance is NIID is the most common problem studied in recent research, as this NIID condition is important for monitoring production processes by means of traditional control charts, such as X and $\bar{X}$ charts.

SVM is used as the recognition algorithm in the proposed CCPR system. Adopting this algorithm, authors such as Chinnam (2002) and Xanthopoulos and Razzaghi (2014) achieved good pattern recognition accuracies. Wu et al. (2015) developed a binary tree SVM-based recognition system, also achieving good recognition accuracies. Khormali and Addeh (2016) described a CCPR system where the SVM is enhanced by using a type-2 fuzzy c-means algorithm. Other authors such as Lu et al. (2011), Du et al. (2013), Xie et al. (2013) and Chompu-Inwai and Thaiupathump (2015) utilised signal processing techniques such as Independent Component Analysis and Wavelet Transforms to pre-process the control chart data, also with good accuracies.

Barghash and Santarisi (2004) studied the effect of training parameters on the performance of the CCPR system, finding that the values range used during pattern generation greatly affects Type 1 and Type 2 errors. Cheng and Cheng (2008) highlighted the importance of the same parameter in the generalisability of the CCPR model.

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

Guh and Tannock (1999) were the first authors that attempted to obtain relevant information about the pattern once it had been identified as one of the seven simple types. They designed an ensemble of subsystems where, in a first step, the pattern is recognised and, in the second step, the magnitude of the pattern is obtained. The creation of ensembles of subsystems for pattern identification and posterior magnitude identification seems to be the most common approach in the work reviewed. For example, see Guh (2003, 2005), Jiang et al. (2009) and Shaban and Shalaby (2012).

2.3. CCPR for autocorrelated processes

The identification of patterns where the inherent noise is not NIID is an infrequently studied issue, so few papers dealing with this topic were found. The task of CCPR for autocorrelated processes can be divided into two: identifying changes in process mean, and identifying abnormal patterns like those proposed in Western Electric Company (1958). Furthermore, all CCPR systems developed so far have assumed that the inherent disturbance can be represented by an AR(1) model. CCPR systems for disturbances modelled by other types of times series such as MA(1) and ARMA(1,1) are needed.

The first attempts to apply ML algorithms for CCPR with autocorrelated processes were intended to detect only mean shifts. Chiu et al. (2001) utilised a back-propagation neural network (BPNN) to identify mean shifts in AR(1) processes with varying autocorrelation levels. Hwarng (2004) monitored the mean value of an autocorrelated process by means of a BPNN, comparing the monitoring capability with those achieved using Exponentially Weighted Moving Average (EWMA), special cause control and other charts. Zobel et al. (2004) developed a BPNN-based technique for CCPR for recognising process mean shifts, incorporating a data processing classification algorithm. Hwarng (2005) proposed a neural-network-based identification system for detecting mean shifts and correlation changes. Wu

For Peer Review Only

and Yu (2010) used a selective network ensemble approach called Discrete Particle Swarm Optimisation for detecting both mean and variance shifts. Guh (2008) developed a real-time recognition ensemble with nineteen on-line Learning Vector Quantisation (LVQ) networks for identifying the simple patterns in control charts, training one on-line LVQ for each correlation level. Cheng and Cheng (2008) developed a neural network recogniser for autocorrelated processes, using Haar Discrete Wavelet Transform (HDW) to denoise, decorrelate and extract features for the CCPR system.

Noorossana et al. (2003) were the first to apply neural networks for detecting and classifying non-random disturbances. The patterns under study were referred to as *level shifts, additive outliers* and *innovation outliers.*

Most recent authors have focused on recognition of simple patterns by on-line ensembles. Lin et al. (2011) designed an on-line real-time CCPR using SVMs as pattern classifiers, training one SVM for each autocorrelation level. Yang and Zhou (2015) developed an LVQ-based ensemble for on-line pattern recognition of the seven simple patterns, providing the autocorrelation level as additional information about the processes.

Cheng and Cheng (2008), Lin et al. (2011) and Yang and Zhou (2015) used an estimator of autocorrelation for common-cause charts in their recognition systems. They neglected the biasing effect on the autocorrelation coefficient caused by the abnormal disturbance (Woodall and Faltin 1993, Dyer et al. 2003), so training the ML algorithms based on erroneous autocorrelation coefficients.

This paper proposes a methodology for generating training patterns for CCPR systems taking into account the possible existence of assignable causes. A recognition model will be presented for each of the following stationary time series models, AR(1), MA(1) and ARMA(1,1). The models are independent of the autocorrelation level.

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

3 CCPR for autocorrelated processes and parameter estimation

3.1. PGS for autocorrelated processes

Synthetic patterns generated from PGSs for the purpose of training CCPR systems should fulfil the following conditions:

- Being generated from a broad range of fully randomised training pattern parameters (slope, shift magnitude, break point position, systematic departure (Hachicha & Ghorbel 2012), amplitude and frequency of cycles). This is to ensure that the CCPR system is able to identify a wide variety of patterns.
- Having parameters that are statistically significant. This is to ensure that correct decision boundaries for pattern classification are obtained by the ML algorithm. For example, the slope of an UT pattern could be reduced or completely removed by noise. The statistical significance of the slope must be tested and found to be positive for a pattern to be classified as a Trend. If this condition is fulfilled, correct decision boundaries for pattern classification can be obtained by the ML algorithm.
- Being generated by an automatic and reproducible procedure, so allowing other researchers to generate the same patterns for comparison purposes.

In order to produce synthetic patterns that fulfil these conditions, the proposed PGS comprises the following steps:

i. Generate one of the seven simple patterns using expressions A5 - A9 from the Appendix and randomly choosing the magnitude β of the pattern.

Thus, the control chart data can be modelled by a NLM-ARMA of the following type:

$$Y_t = D_t + N_t \quad (2)$$

For Peer Review Only

where D_t represents the magnitude of the abnormal pattern and is given by a nonlinear model (to be defined in later steps), and N_t represents the inherent disturbance to be modelled by one of the three stationary processes utilised in this work (see Appendix).

ii. Determine the type of mean change in the pattern. Following the methodology proposed by De La Torre Gutierrez and Pham (2016), the determination of the type of mean change is considered as a problem of selecting between two models, in this case, two NLM-ARMA models when Y_t is as shown in equation (3). The most likely break point is found by fitting piecewise NLM-ARMA models at time $\tau = 16, 17, \dots, (n-15)$ (see Appendix).

$$Y_t = \beta_0 + \beta_1 t + \beta_2 d + \beta_3 sin\left(\frac{2\pi t}{\beta_4}\right) + \beta_5(-1)^t + N_t \quad (3)$$

where the β_1 to β_5 and *d* represent the pattern parameters as defined in the Appendix. β_0 and N_t represent the intercept with the y-axis of the regression model and the random error at time *t*.

The Bayesian Information Criterion (BIC) is extracted from each fitted model and the one with the minimum BIC value is selected. Such a selected model represents the model with the most likely break point.

iii. Compare the model representing the most likely break point with the one where no break points are assumed. Equation (4) represents the model to be used for Y_t when there is no break point. Models (3) and (4) represent two NLM-ARMA nested models; the full model is that with the most likely break point and the reduced model the one that does not consider the existence of break points.

$$Y_t = \beta_0 + \beta_1 t + \beta_3 sin\left(\frac{2\pi t}{\beta_4}\right) + \beta_5(-1)^t + N_t \quad (4)$$

The hypotheses for this problem of model selection are:

H_0: There are no break points (the reduced model fits better)

For Peer Review Only

H_1: A break point is detected (the full model fits better)

The F-statistic shown in (5) is used to determine which model better fits to the control data.

$$F_{v_2}^{v_1} = (SSE_{full} - SSE_{reduced})/(SSE_{full}/(n - k - 1)) \quad (5)$$

where SSE_{full} and $SSE_{reduced}$ represent the Sum of Squared Error (SSE) of the full and the reduced model, respectively. F represents an F-distribution with one degree of freedom in the numerator ($\upsilon_1 = 1$) and n – k - 1 degrees of freedom in the denominator ($\upsilon_2 = n - k - 1$), k being the number of parameters in the full model.

iv. Determine pattern class. Once it has been decided whether (3) or (4) fits the pattern better, the pattern class is determined by the corresponding β value that is statistically significant. Three significance levels were utilised α=0.01, 0.02 and 0.03 in this work.

Figure 2 summarises the proposed PGS.

Figure 2: Flowchart of the proposed PGS

3.2. Training of the ML

Three sets of 5,600 training patterns, 800 of each type, were generated using α = 0.01, 0.02, 0.03 using the proposed PGS and another set was created conventionally (Pham et al. 2006, Pham and Oztemel 1996).

The Bees Algorithm proposed by Pham et al. (2006) was used to find the best sets of free parameters that ensure the minimum misclassification rate with the 5-fold cross-validation technique.

The Bees Algorithm is a novel population-based optimisation algorithm inspired by the foraging activities of bees. The algorithm was chosen due to its simplicity and proven search

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

and optimisation capabilities (Pham and Castellani 2014). The parameters of the Bees Algorithm used in this work are shown below.

Initial population (*n*)	30
Number of "best" sites (*m*)	5
Number of "elite" sites (*e*)	2
Patch size for Cost (C) parameter (*ngh-c)*	0.5
Patch size for Kernel parameters (*ngh-k)*	0.02
Number of elite bees for the elite sites (*ne)*	4
Number of bees for the remaining "best" points (*nb)*	2

For more explanation of the algorithm and parameters, see Pham et al. (2006) and Pham and Castellani (2014).

Once the best sets of free parameters are found and the SVMs are trained, the pattern recognition accuracies of the recognition systems can be assessed.

3.3. Determination of the best arrangement of input factors

For testing purposes, one hundred sets of 700 patterns, 100 of each type, were generated using the best arrangement, and the pattern recognition accuracies obtained.

The three factors studied in the first step of the analysis are as follows:

- IRT: Standardised raw data and shape features. To standardise the raw data, the control chart data are rescaled using the following expression:

$$Z_t = \frac{Y_t - \bar{Y}}{S_y} \qquad (6)$$

where Z_t represents the scaled variable to be used as input for the training of the SVM, $\bar{Y}$ is the estimated mean value of the current control chart and S_y is the estimated standard deviation.

The shape features utilised here are those initially proposed by Pham and Wani (1997) and improved by Gauri (2010). These features have advantages in CCPR problems

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

where the process to be monitored is NIID. These are reduction in the training time, increase in pattern recognition accuracy and independence from the data scale. The performance of these features has not been assessed in CCPR systems where the inherent disturbance is not NIID. Therefore, this is a factor to be considered during the first step of the analysis.

- PGS. The PGS adopted in this work deals with the correct categorisation of training patterns before they are input to the CCPR system. The performance of the recognition systems trained using patterns generated at three α levels is assessed, and compared with the recognition achieved without the PGS.
- SVM kernel. Four kernels were tested: Radial Basis Function (RBF), Laplace (LAPLA), Hyperbolic Tangent (TANH) and Bessel (BESS).

4 Results

4.1. Analysis of the PGS

Three sets of 70,000 random vectors, 10,000 of each pattern type, were initially generated. Each vector $(Y_1, Y_2, \dots, Y_{60})$ represents a quality characteristic sampled at time $t_1, t_2, \dots, t_{60}$. Each set of random vectors was created using one of the three first-order stationary models. The patterns were passed through the PGS. Three significance levels were set, namely, α = 0.01, 0.02 and 0.03. The allocation of the patterns to the different classes is shown in Table 1.

In Table 1, for each of the models, the column "Retained" gives the percentages of patterns for which the final classification by the PGS agrees with the classification when the patterns were initially produced in step (i). The column "Reclassified" shows the percentages of patterns for which the classification was changed after they were passed through the PGS.

For Peer Review Only

The column "Discarded" gives the percentages of patterns rejected by the PGS as not recognisable due to the low statistical significance of the parameters that characterise them.

Table 1: Allocation of patterns passed through the proposed PGS (%)

It can also be observed that when $\alpha = 0.01$, 24.26%, 32.91% and 43.40% of the patterns were discarded for AR, MA and ARMA processes, respectively. When this α level was used, the pattern with the highest reclassification rate was the Normal pattern for the MA and ARMA models, and the US pattern for the AR model. This was probably due to the initial blend of inherent noise and abnormal disturbance. With the significance level set to $\alpha = 0.02$, 35.96%, 44.09% and 53.89% of the patterns were discarded when the AR, MA and ARMA models were used, respectively. In the case of $\alpha = 0.03$, the percentage of discarded patterns increased to 44.43%, 51.93% and 60.78% for the aforementioned three processes.

4.2. Overall accuracies

To simplify the analysis of the results, it will be carried out in two steps. First, the best arrangement amongst the input factors for each disturbance model is determined; these factors are IRT, PGS and kernel of the recognition system. Using the best arrangement determined in the first step of the analysis, the performance of the CCPR system for a set of patterns generated with a specific autocorrelation level is then assessed.

Figures 3, 4 and 5 show the accuracies found for the three inherent noise models. These accuracies are disaggregated by IRT, Kernel and pattern type, the PGS being the constant factor in each plot. In these plots, it can be observed that the three lines corresponding to the three α levels are nearly coincident; this indicates that the proposed PGS is robust to changes in α values.

Figure 3: Accuracies achieved from AR process (%)

For Peer Review Only

Figure 4: Accuracies achieved from MA process (%)

Figure 5: Accuracies achieved from ARMA process (%)

ANOVA of type 4x2x4 with single, double and triple interactions was utilised to determine which factors were significant for the pattern recognition accuracies. The p-values of each factor are shown in Table 2.

Table 2: P-values of the of the three factors obtained from the ANOVA

Tukey post-hoc tests were used to determine the best arrangement for each disturbance model; results are summarised in Table 3.

Table 3: Best arrangement for AR, MA and ARMA processes

4.3. Analysis of the best arrangement of input factors

To assess the performance of the SVM trained with the best combination of factors, one hundred sets of 700 patterns, 100 of each type were generated for each autocorrelation level. The combination used is: PGS with $\alpha = 0.01$, Raw data and RBF kernel.

Table 4 gives the accuracies found for the AR and ARMA models disaggregated by pattern type and autocorrelation level. For the case of the AR process, an overall accuracy of 90.03% was found. The US pattern type was the one with the least recognition accuracy. It was also found that the accuracy for all pattern types was lowest with $\phi \geq 0.70$.

The accuracies achieved for the MA and ARMA processes are shown in Table 5, disaggregated by pattern type and moving-average level. For the MA process, the US pattern

For Peer Review Only

class and patterns of all types generated with $\theta \geq 0.70$ yielded the lowest recognition accuracies.

To obtain the accuracies for the ARMA process of Table 4, the ϕ level was fixed and the θ values were randomised. It can be observed that the overall accuracy for the ARMA process considering all the ϕ levels is 89.08%, the Normal pattern having the lowest accuracy. Again, an autocorrelation level of $\phi \geq 0.70$ produced the least accuracy.

Likewise in Table 5, to obtain the accuracies for the ARMA process, the θ was fixed and the ϕ levels were randomised. It can be seen that the lowest accuracies were obtained for Normal patterns and $\theta \geq 0.70$.

Table 4: Accuracies for AR and ARMA processes by ϕ values (%)

Table 5: Accuracies for MA and ARMA processes by θ values (%)

The performance of the proposed CCPR systems was also analysed by type and magnitude of the pattern. Table 6 presents the accuracies obtained by the CCPR systems of the three process types disaggregated by pattern type and magnitude. It can be observed that, for the three process types, the US and DS patterns showed the lowest and the highest accuracies, respectively. It can also be seen that the recognition accuracies decreased as the pattern magnitude was reduced.

Table 6: Accuracies for the three processes by pattern magnitude (%)

For Peer Review Only

4.4. Real data application

Figure 6: Thickness of metallic film in the early stages of the development of an electronic device (Box et al. 2009)

To demonstrate the ability of the proposed CCPR system to handle real data, thickness measurements for a very thin metallic film in the early stages of the development of an electronic device data taken from Box et al. (2009) were utilised (see Figure 6). Box et al. (2009) highlighted the existence of an assignable cause that abruptly increased the metallic film thickness after 30 observations.

The CCPR trained with the best arrangement of input factors for the AR model was applied, and an US pattern was identified.

In order further to categorise the pattern recognised by the aforementioned CCPR system, the two NLM-ARMA models proposed in this paper were fitted to the original data. The p-value corresponding to the F-test for nested models was found to be 0.000002; therefore, the full model fitted better the data. The most likely breakpoint was detected at $\tau = 30$.

Table 7 shows all the values for the fitted full model. It can be observed that the parameter related to the Shift pattern is statistically significant and greater than zero. Thus like the CCPR, the NLM-ARMA classifies the pattern as an US pattern.

Table 7: ANOVA of the NLM-ARMA model fitted to the metallic film thickness data

For Peer Review Only

5 Conclusion

The literature review uncovered very few papers dealing with monitoring of non-NIID production processes, the AR(1) model being the only model studied. First-order stationary models are used to represent many continuous production processes. It is necessary to develop robust CCPR systems that allow the automated identification of abnormal patterns where these types of processes are monitored using control charts.

The proposed PGS fulfils three conditions: generality, comparability and facilitation of the extraction of further information from the pattern. Generality was achieved by employing a broad range of pattern parameters during the initial pattern generation and the total randomisation of parameters, including the break point position and the amplitude of Cyclic patterns. Comparability between different studies is possible as the proposed PGS is a standard technique for producing the same data for training CCPR systems, requiring only the significance level to be set. Further information (for example, the amplitude of cycles) regarding the abnormal pattern identified can be extracted as the decision boundaries were estimated using patterns for which the statistical significance of the parameters was tested. Therefore, the CCPR system can categorise the presented patterns into the class for which the related parameter is the most significant when a NLM-ARMA model is fitted. Furthermore, by fitting this dynamic model to the control chart data, the PGS is able to divide signals into inherent noise and fault signal.

From the first part of the recognition accuracy analysis, of the four different PGS ($\alpha = 0.01$, 0.02, 0.03 and conventional PGS), the proposed PGS significantly increased the pattern recognition accuracy for the three models used. The three α levels achieved nearly the same results, this indicates that the proposed PGS is robust to changes in α values. It is

For Peer Review Only

recommended to use low significance levels such as 0.01 in order to reduce computational efforts and generate fewer patterns.

Regarding the IRT, the shape features and raw data achieved similar accuracies in the MA and ARMA models, but lower accuracies were obtained with shape features when the AR process was utilised. Two kernels, RBF and Bessel, showed the greatest recognition accuracies.

In the second part of the analysis, it was found that the values of θ and ϕ greatly affected the pattern recognition accuracy. For the MA and ARMA processes, the lowest accuracy was found when $\theta \geq 0.70$, and the highest accuracy was achieved when $-0.5 < \theta \leq -0.3$. For AR and ARMA processes, the same behaviour was observed, with the lowest accuracy obtained when $\phi \geq 0.70$, and the highest accuracy when $-0.5 < \phi \leq -0.3$. Regarding the patterns, the Normal and US types were those that gave the lowest accuracies.

The implementation of the proposed recognition system on a real production line is straightforward; the operator only needs to set the observation window to 60, collect data from the production variable being monitored and then input it to a CCPR system trained using the best arrangement of factors determined in section 4.2.

Future research can proceed in three directions: (i) developing CCPR systems and PGS for non-stationary processes such as IMA and ARIMA, (ii) expanding the proposed methodology to patterns combining two or more pattern types, and (iii) developing PGS and CCPR systems for multivariate and stationary variables, using different control chart types to monitor quality characteristics.

For Peer Review Only

6 Acknowledgements

The first author acknowledges the support of the Mexican National Council for Science and Technology, CONACYT, enabling him to pursue graduate studies at University of Birmingham – United Kingdom, under grant number 381672. The authors also acknowledge The Birmingham Environment for Academic Research (BEAR) project for their technical support.

For Peer Review Only

7 References

Alwan, L.C., 1991. Autocorrelation: Fixed versus variable control limits. *Quality engineering*, 4(2), pp.167–188.

Barghash, M.A. & Santarisi, N.S., 2004. Pattern recognition of control charts using artificial neural networks—analyzing the effect of the training parameters. *Journal of Intelligent Manufacturing*, 15(5), pp.635–644.

Box, G.E.P., Luceño, A. & Paniagua-Quiñones, M. del C., 2009. *Statistical control by monitoring and adjustment*, John Wiley & Sons.

Box, G.E.P., Luceño, A. & Paniagua-Quiñones, M.D.C., 2011. *Statistical Control by Monitoring and Adjustment: Second Edition*,

Boyles, R.A., 2000. Phase I Analysis for Autocorrelated Processes. *Journal of Quality Technology*, 32(4), pp.395–409.

Cheng, H.P. & Cheng, C.S., 2008. Denoising and feature extraction for control chart pattern recognition in autocorrelated processes. *International Journal of Signal and Imaging Systems Engineering*, 1(2), pp.115–126.

Chinnam, R.B., 2002. Support vector machines for recognizing shifts in correlated and other manufacturing processes. *International Journal of Production Research*, 40(17), pp.4449–4466.

Chiu, C.C., Chen, M.K. & Lee, K.M., 2001. Shifts recognition in correlated process data using a neural network. *International Journal of Systems Science*, 32(2), pp.137–143.

Chompu-Inwai, R. & Thaiupathump, T., 2015. Improved ICA-based mixture control chart patterns recognition using shape related features. In *2015 IEEE Conference on Control and Applications, CCA 2015 - Proceedings*. IEEE, pp. 484–489.

Du, S.C., Huang, D.L. & Lv, J., 2013. Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines. *Computers & Industrial Engineering*, 66(4), pp.683–695.

Dyer, J.N., Adams, B.M. & Conerly, M.D., 2003. The Reverse Moving Average Control Chart for Monitoring Autocorrelated Processes. *Journal of Quality Technology*, 35(2), pp.139–152.

Gauri, S.K., 2010. Control chart pattern recognition using feature-based learning vector quantization. *International Journal of Advanced Manufacturing Technology*, 48(3), pp.1061–1073.

Guh, R.S., 2005. A hybrid learning-based model for on-line detection and analysis of control chart patterns. *Computers & Industrial Engineering*, 49(1), pp.35–62.

Guh, R.S., 2003. Integrating artificial intelligence into on-line statistical process control. *Quality and Reliability Engineering International*, 19(1), pp.1–20.

Guh, R.S., 2008. Real-time recognition of control chart patterns in autocorrelated processes using a learning vector quantization network-based approach. *International Journal of Production Research*, 46(14), pp.3959–3991.

For Peer Review Only

Guh, R.S. & Tannock, J.D.T., 1999. A neural network approach to characterize pattern parameters in process control charts. *Journal of Intelligent Manufacturing*, 10(5), pp.449–462.

Hachicha, W. & Ghorbel, A., 2012. A survey of control-chart pattern-recognition literature (1991-2010) based on a new conceptual classification scheme. *Computers and Industrial Engineering*, 63(1), pp.204–222.

Hwarng, H.B., 2004. Detecting process mean shift in the presence of autocorrelation: a neural-network based monitoring scheme. *International Journal of Production Research*, 42(3), pp.573–595.

Hwarng, H.B., 2005. Simultaneous identification of mean shift and correlation change in AR(1) processes. *International Journal of Production Research*, 43(9), pp.1761–1783.

Hyndman, R.J. & Athanasopoulos, G., 2014. *Forecasting: principles and practice*, OTexts.

Jiang, P., Liu, D. & Zeng, Z., 2009. Recognizing control chart patterns with neural network and numerical fitting. *Journal of Intelligent Manufacturing*, 20(6), pp.625–635.

Jiang, W. & Tsui, K.-L., 2002. SPC Monitoring of MMSE- and PI-Controlled Processes. *Journal of Quality Technology*, 34(4), pp.384–398.

Kazemi, M.S. et al., 2015. A hybrid method for estimating the process change point using support vector machine and fuzzy statistical clustering. *Applied Soft Computing*, 40, pp.507–516.

Khormali, A. & Addeh, J., 2016. A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine. *ISA Transactions*, 63, pp.256–264.

De La Torre Gutierrez, H. & Pham, D.T., 2016. Estimation and generation of training patterns for control chart pattern recognition. *Computers and Industrial Engineering*, 95, pp.72–82.

Lesany, S.A., Koochakzadeh, A. & Fatemi Ghomi, S.M.T., 2013. Recognition and classification of single and concurrent unnatural patterns in control charts via neural networks and fitted line of samples. *International Journal of Production Research*, 52(6), pp.1771–1786.

Lin, S.-Y., Guh, R.-S. & Shiue, Y.-R., 2011. Effective recognition of control chart patterns in autocorrelated data using a support vector machine based approach. *Computers & Industrial Engineering*, 61(4), pp.1123–1134.

Longnecker, M.T. & Ryan, T.P., 1992. Charting correlated process data. *Texas A&M University Department of Statistics Technical Report*, 166.

Lu, C.-J., Shao, Y.E. & Li, P.-H., 2011. Mixture control chart patterns recognition using independent component analysis and support vector machine. *Neurocomputing*, 74(11), pp.1908–1914.

Lu, C.-W. & Reynolds Jr., M.R., 1999. Control Charts for Monitoring the Mean and Variance of Autocorrelated Processes. *Journal of Quality Technology*, 31(3), pp.259–274.

For Peer Review Only

MacGregor, J.F., 1988. On-line statistical process control. *Chemical Engineering Progress*, 84(10), pp.21–31.

Matsumoto, M. & Nishimura, T., 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, 8(1), pp.3–30.

Montgomery, D.C. et al., 2000. Integrating Statistical Process Monitoring. *Quality and Reliability Engineering International*, 16(September 1999), pp.515–525.

Montgomery, D.C., 2009. *Introduction to statistical quality control*.

Nembhard, H.B. & Kao, M.S., 2003. Adaptive Forecast-Based Monitoring for Dynamic Systems. *Technometrics*, 45(3), pp.208–219.

Noorossana, R., Farrokhi, M. & Saghaei, A., 2003. Using Neural Networks to Detect and Classify Out-of-control Signals in Autocorrelated Processes. *Quality and Reliability Engineering International*, 19(6), pp.493–504.

Pankratz, A., 1991. *Forecasting with Dynamic Regression Models*, Hoboken, NJ, USA: John Wiley & Sons, Inc.

Pham, D. T. & Oztemel, E., 1996. *Intelligent quality systems*, Springer Science & Business Media.

Pham, D.T. et al., 2006. Application of the Bees Algorithm to the Training of Learning Vector Quantisation Networks for Control Chart Pattern Recognition. *2006 2nd International Conference on Information & Communication Technologies*, 1, pp.0–5.

Pham, D.T. et al., 2006. The Bees Algorithm - A Novel Tool for Complex Optimisation Problems. In *Intelligent Production Machines and Systems - 2nd I*PROMS Virtual International Conference 3-14 July 2006*. pp. 454–459.

Pham, D.T. & Castellani, M., 2014. Benchmarking and comparison of nature-inspired population-based continuous optimisation algorithms. *Soft Computing*, 18(5), pp.871–903.

Pham, D.T. & Wani, M.A., 1997. Feature-based control chart pattern recognition. *International Journal of Production Research*, 35(7), pp.1875–1890.

Psarakis, S. & Papaleonida, G., 2007. SPC procedures for monitoring autocorrelated processes. *Quality Technology & Quantitative Management*, 4(4), pp.501–540.

Shaban, A. & Shalaby, M.A., 2012. A double neural network approach for the identification and parameter estimation of control chart patterns. *International Journal of Quality Engineering and Technology*, 3(2), p.124.

Tsung, F., Wu, H. & Nair, V.N., 1998. On the Efficiency and Robustness of Discrete Proportional-Integral Control Schemes. *Technometrics*, 40(3), pp.214–222.

Wang, K. & Tsung, F., 2007. Monitoring feedback-controlled processes using adaptive T2 schemes. *International Journal of Production Research*, 45(23), pp.5601–5619.

Western Electric Company, 1956. Statistical Quality Control Handbook.

For Peer Review Only

Woodall, W.H. & Faltin, F.W., 1993. Autocorrelated data and SPC. *ASQC Statistics Division Newsletter*, 13(4), pp.18–21.

Wu, B. & Yu, J.B., 2010. A neural network-based on-line monitoring model of process mean and variance shifts. *Proceedings of the International Conference on E-Business and E-Government, ICEE 2010*, 37(6), pp.2615–2618.

Wu, C., Liu, F. & Zhu, B., 2015. Control chart pattern recognition using an integrated model based on binary-tree support vector machine. *International Journal of Production Research*, 53(7), pp.2026–2040.

Xanthopoulos, P. & Razzaghi, T., 2014. A weighted support vector machine method for control chart pattern recognition. *Computers & Industrial Engineering*, 70, pp.134–149.

Xie, L. et al., 2013. Concurrent control chart patterns recognition with singular spectrum analysis and support vector machine. *Computers & Industrial Engineering*, 64(1), pp.280–289.

Yang, W.A. & Zhou, W., 2015. Autoregressive coefficient-invariant control chart pattern recognition in autocorrelated manufacturing processes using neural network ensemble. *Journal of Intelligent Manufacturing*, 26(6), pp.1161–1180.

Zhang, N.F., 1997. Detection capability of residual control chart for stationary process data. *Journal of Applied Statistics*, 24(4), pp.475–492.

Zhang, N.F. & Pollard, J.F., 1994. Analysis of Autocorrelations Processes in Dynamic. *Technometrics*, 36(4), pp.354–368.

Zobel, C.W., Cook, D.F. & Nottingham, Q.J., 2004. An augmented neural network classification approach to detecting mean shifts in correlated manufacturing process parameters. *International Journal of Production Research*, 42(4), pp.741–758.

For Peer Review Only

8 Appendix

The first step in producing patterns for the proposed PGS is to generate the inherent noise according to one of the three models. To generate such an inherent disturbance, it is necessary to create a white-noise vector. Following the methodology proposed by De La Torre Gutierrez and Pham (2016), an NIID vector, e_t, is produced based on the pseudo-random number generator proposed by Matsumoto and Nishimura (1998), i.e.

$$e_t = Normal(\mu = 0, \sigma_e = 1) \quad \text{(A1)}$$

The methodology for generating ARMA (1,1) processes from normal white-noise proposed by Box et al. (2011) was adopted in this work.

Therefore, the inherent noise represented as an ARMA(1,1) process is represented as:

$$N_t = \phi N_{t-1} - \theta e_{t-1} + e_t \quad \text{(A2)}$$

where if $\phi = 0$, an MA(1) process is obtained; and if $\theta = 0$, an AR(1) model is obtained. Both ϕ and θ are in the range (-1, +1).

The standard deviation of this ARMA(1,1) process can be estimated by means of the following expression:

$$\sigma_N = \sqrt{\frac{1 + \theta^2 - 2\phi\theta}{1 - \phi^2}}\sigma_e \quad \text{(A3)}$$

The model for disturbances in control charts, assuming independence between the inherent noise and deterministic disturbance, is the following:

$$Y_t = D_t + N_t \quad \text{(A4)}$$

where Y_t represents the process to be monitored, D_t the deterministic disturbance and N_t the inherent noise. The seven simple patterns can be generated by using the following expressions:

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

- Normal (NORM):

$$Y_t = \mu + N_t \quad \text{(A5)}$$

- Upward/Downward Trend (UT/DT):

$$Y_t = \mu + N_t \pm \beta_1 t \quad \text{(A6)}$$

- Upward/Downward Shift (US/DS):

$$Y_t = \mu + N_t \pm \beta_2 d \quad \text{(A7)}$$

- Cyclic (CYC):

$$Y_t = \mu + N_t + \beta_3 \, sin\left(\frac{2\pi t}{\beta_4}\right) \quad \text{(A8)}$$

- Systematic (SYS):

$$Y_t = \mu + N_t + \beta_5 (-1)^t \quad \text{(A9)}$$

Without loss of generality, in this work μ is set to zero.

The meaning and range of the pattern parameters used are as described in De La Torre Gutierrez and Pham (2016).

The magnitude shift, slope, systematic departure and cyclic amplitude were kept within in $6\sigma_N$ control limits in the inspection window. The frequency of the cyclic pattern was determined to show at least four cycles in the inspection window.

The break point position was randomly chosen between τ =16 and τ = n-15. This is due the number of parameters and degrees of freedom available during the estimation of the parameters when NLM-ARMA is utilised.

http://mc.manuscriptcentral.com/tprs Email: ijpr@tandf.co.uk

For Peer Review Only

TABLES

Table 1: Allocation of patterns passed through the proposed PGS (%)

	AR			MA			ARMA		
	Retained	Reclassified	Discarded	Retained	Reclassified	Discarded	Retained	Reclassified	Discarded
α = 0.01	64.45	11.29	24.26	57.12	9.97	32.91	46.67	9.92	43.40
NORM	62.70	20.53	16.77	52.17	22.94	24.89	38.08	26.51	35.40
UT	67.57	2.60	29.83	58.79	1.80	39.41	46.25	3.17	50.58
DT	67.65	2.83	29.52	58.78	1.86	39.37	46.16	3.18	50.66
US	59.78	21.43	18.80	57.36	17.85	24.79	54.80	13.62	31.58
DS	59.89	21.27	18.85	57.38	18.10	24.52	54.78	13.46	31.76
CYC	69.63	1.49	28.88	56.94	2.14	40.92	41.95	2.15	55.90
SYS	63.96	8.88	27.16	58.41	5.10	36.49	44.71	7.36	47.93
α = 0.02	54.30	9.74	35.96	47.34	8.57	44.09	37.83	8.28	53.89
NORM	48.20	25.52	26.28	38.74	26.16	35.10	26.98	27.29	45.73
UT	53.55	2.51	43.94	45.87	1.76	52.37	34.85	2.72	62.43
DT	53.77	2.68	43.55	45.98	1.81	52.21	34.87	2.77	62.36
US	58.12	14.74	27.15	54.47	12.33	33.20	51.23	8.87	39.90
DS	58.31	14.58	27.11	54.47	12.64	32.89	51.17	8.86	39.96
CYC	56.28	0.96	42.76	45.34	1.39	53.28	31.69	1.47	66.84
SYS	51.90	7.17	40.93	46.52	3.89	49.59	34.01	5.98	60.02
α = 0.03	46.85	8.72	44.43	40.48	7.59	51.93	32.01	7.21	60.78
NORM	38.58	27.48	33.94	30.35	26.81	42.84	20.54	26.43	53.02
UT	43.77	2.39	53.84	37.14	1.69	61.18	27.72	2.42	69.87
DT	43.89	2.53	53.58	37.35	1.76	60.88	27.72	2.49	69.79
US	55.78	11.02	33.20	51.55	9.28	39.18	48.00	6.39	45.61
DS	55.93	10.84	33.23	51.41	9.38	39.21	47.93	6.47	45.60
CYC	46.79	0.69	52.52	37.28	0.99	61.73	25.06	1.12	73.82
SYS	43.23	6.09	50.68	38.25	3.22	58.53	27.10	5.16	67.74

For Peer Review Only

Table 2: P-values of the of the three factors obtained from the ANOVA

p-values			
Factor	AR	MA	ARMA
PGS	0.0000	0.0000	0.0000
KERNEL	0.0000	0.0000	0.0000
IRT	0.0000	0.5150	0.0638
PGS*KERNEL	0.0186	0.1800	0.1510
KERNEL*IRT	0.0000	0.0000	0.0000
PGS*IRT	0.0583	0.6680	0.2197
PGS*KERNEL*IRT	0.0000	0.0000	0.0000

For Peer Review Only

Table 3: Best arrangement for AR, MA and ARMA processes

Disturbance model	AR	MA	ARMA
PGS	$\alpha = 0.01$, 0.02, 0.03	$\alpha = 0.01$, 0.02, 0.03	$\alpha = 0.01$, 0.02
IRT	Raw data	Raw data, Shape features	Raw data, Shape features
Kernel	Bessel/RBF	Bessel/RBF	Bessel/RBF

For Peer Review Only

Table 4: Accuracies for AR and ARMA processes by ϕ values (%)

ϕ	OVERALL	NORM	UT	DT	US	DS	CYC	SYS
AR	90.03	83.92	91.08	93.60	79.70	96.64	90.97	94.28
$\phi \leq -0.7$	91.86	77.50	91.49	94.21	93.17	99.82	91.80	95.00
$-0.7 < \phi \leq -0.5$	92.97	80.00	93.51	93.65	96.54	99.78	93.13	94.33
$-0.5 < \phi \leq -0.3$	92.98	82.63	93.93	93.46	94.92	98.52	93.46	93.83
$-0.3 < \phi \leq 0.0$	92.19	87.17	93.07	93.07	87.94	97.44	92.51	94.12
$0.0 < \phi \leq 0.3$	90.47	89.52	91.94	93.20	77.37	95.46	91.69	94.10
$0.3 < \phi \leq 0.5$	89.22	90.24	91.29	93.79	68.76	95.06	91.08	94.31
$0.5 < \phi < 0.7$	87.39	86.13	89.15	93.71	65.27	94.54	88.99	93.91
≥ 0.7	83.15	78.15	84.26	93.68	53.65	92.53	85.09	94.66
ARMA	89.08	80.61	89.63	93.17	82.43	95.55	89.44	92.77
$\phi \leq -0.7$	88.22	67.83	88.72	90.76	90.94	99.03	87.40	92.83
$-0.7 < \phi \leq -0.5$	90.26	75.62	89.78	92.16	93.68	98.56	90.33	91.67
$-0.5 < \phi \leq -0.3$	91.75	82.37	92.35	93.95	93.02	97.13	91.26	92.16
$-0.3 < \phi \leq 0.0$	91.64	86.73	92.05	94.08	88.27	96.53	91.63	92.17
$0.0 < \phi \leq 0.3$	91.22	88.64	91.79	94.68	83.17	95.77	91.68	92.81
$0.3 < \phi \leq 0.5$	90.55	88.27	91.74	94.49	79.30	94.97	91.17	93.93
$0.5 < \phi < 0.7$	88.00	84.72	88.42	93.98	72.63	93.83	88.66	93.73
$\phi \geq 0.7$	81.05	70.70	82.15	91.28	58.39	88.60	83.36	92.85

For Peer Review Only

Table 5: Accuracies for MA and ARMA processes by θ values (%)

θ	OVERALL	NORM	UT	DT	US	DS	CYC	SYS
MA	91.47	84.86	92.92	94.39	82.66	97.75	93.15	94.59
$\theta \leq -0.7$	91.88	80.39	89.36	94.24	95.90	99.99	88.76	94.52
$-0.7 < \theta \leq -0.5$	92.56	80.75	92.40	93.82	94.15	99.97	92.32	94.54
$-0.5 < \theta \leq -0.3$	93.22	81.64	94.59	94.16	93.18	99.96	94.39	94.60
$-0.3 < \theta \leq 0.0$	92.86	83.07	95.20	94.15	88.41	99.32	95.41	94.43
$0.0 < \theta \leq 0.3$	92.18	86.99	95.13	94.29	80.71	98.14	95.29	94.73
$0.3 < \theta \leq 0.5$	90.41	87.48	93.54	94.84	72.76	96.06	93.65	94.51
$0.5 < \theta < 0.7$	89.85	89.27	92.99	95.07	68.99	94.53	93.33	94.79
$\theta \geq 0.7$	88.83	89.29	90.18	94.52	67.21	94.03	92.03	94.57
ARMA	89.48	81.47	89.98	93.42	83.65	95.71	89.50	92.65
$\theta \leq -0.7$	89.21	74.55	87.04	91.52	96.34	98.40	85.57	91.07
$-0.7 < \theta \leq -0.5$	90.23	77.16	89.67	93.29	93.99	98.13	87.87	91.47
$-0.5 < \theta \leq -0.3$	90.84	80.42	91.20	93.74	89.77	97.55	91.00	92.18
$-0.3 < \theta \leq 0.0$	90.71	80.43	92.67	94.59	85.38	96.66	92.61	92.61
$0.0 < \theta \leq 0.3$	90.34	83.97	92.64	94.07	80.50	96.20	91.66	93.35
$0.3 < \theta \leq 0.5$	89.58	84.75	91.04	94.24	76.37	94.78	91.33	94.56
$0.5 < \theta < 0.7$	88.29	85.62	89.30	93.41	74.22	93.68	89.16	92.65
$\theta \geq 0.7$	86.67	84.84	86.29	92.47	72.66	90.29	86.81	93.30

For Peer Review Only

Table 6: Accuracies for the three processes by pattern magnitude (%)

	Magnitude	ARMA	MA	AR
UT	TOTAL	89.80	92.92	91.08
	$0.25 \leq \beta_1 \leq 0.30$	99.74	100.00	100.00
	$0.20 \leq \beta_1 < 0.25$	99.59	100.00	100.00
	$0.15 \leq \beta_1 < 0.20$	99.25	100.00	99.99
	$0.10 \leq \beta_1 < 0.15$	98.47	99.96	99.72
	$0.05 \leq \beta_1 < 0.10$	88.74	95.77	93.43
	$0.00 < \beta_1 < 0.05$	53.01	61.77	53.35
DT	TOTAL	93.30	94.39	93.60
	$-0.25 \geq \beta_1 \geq -0.30$	99.83	100.00	100.00
	$-0.20 \geq \beta_1 > -0.25$	99.75	100.00	100.00
	$-0.15 \geq \beta_1 > -0.20$	99.38	100.00	99.95
	$-0.10 \geq \beta_1 > -0.15$	98.51	99.99	99.62
	$-0.05 \geq \beta_1 > -0.10$	91.52	95.17	92.26
	$0.00 > \beta_1 > -0.05$	70.82	71.17	69.75
US	TOTAL	83.00	82.66	79.70
	$2.5 \leq \beta_2 \leq 3.0$	96.65	98.64	97.00
	$2.0 \leq \beta_2 < 2.5$	96.80	98.63	96.38
	$1.5 \leq \beta_2 < 2.0$	96.34	98.25	95.07
	$1.0 \leq \beta_2 < 1.5$	92.70	96.10	89.58
	$0.5 \leq \beta_2 < 1.0$	72.72	76.87	70.06
	$0.0 < \beta_2 < 0.5$	42.79	27.47	30.09
DS	TOTAL	95.60	97.75	96.64
	$-2.5 \geq \beta_2 \geq -3.0$	97.99	99.96	99.82
	$-2.0 \geq \beta_2 > -2.5$	97.56	98.94	98.79
	$-1.5 \geq \beta_2 > -2.0$	96.12	98.38	97.94
	$-1.0 \geq \beta_2 > -1.5$	95.40	97.96	97.86
	$-0.5 \geq \beta_2 > -1.0$	94.24	96.50	94.93
	$0.0 > \beta_2 > -0.5$	92.31	94.73	90.48
CYC	TOTAL	89.50	93.15	90.97
	$2.5 \leq \beta_3 \leq 3.0$	96.51	98.50	98.76
	$2.0 \leq \beta_3 < 2.5$	94.05	98.41	98.14
	$1.5 \leq \beta_3 < 2.0$	91.27	96.44	96.93
	$1.0 \leq \beta_3 < 1.5$	90.09	93.33	89.37
	$0.5 \leq \beta_3 < 1.0$	79.85	88.20	83.64
	$0.0 < \beta_3 < 0.5$	85.20	84.00	79.00
SYS	TOTAL	92.70	94.59	94.28
	$2.5 \leq \beta_5 \leq 3.0$	99.98	100.00	100.00
	$2.0 \leq \beta_5 < 2.5$	99.98	100.00	100.00
	$1.5 \leq \beta_5 < 2.0$	99.91	100.00	100.00

For Peer Review Only

	$1.0 \leq \beta_5 < 1.5$	99.55	100.00	99.97
	$0.5 \leq \beta_5 < 1.0$	96.51	97.22	96.60
	$0.0 < \beta_5 < 0.5$	60.25	70.33	69.10

For Peer Review Only

Table 7: ANOVA of the NLM-ARMA model fitted to the metallic film thickness data

Parameter	Estimate	Std. Error	t value	p-value
ϕ	-0.2632	0.1075	-2.4484	0.0144
Intercept	80.4461	2.1699	37.0736	0.0000
Slope (Trend)	0.1750	0.0986	1.7748	0.0760
Shift magnitude (Shift)	23.0429	4.0195	5.7328	0.0000
Amplitude (Cyclic)	2.1916	1.4738	1.4870	0.1370
Frequency (Cyclic)	8.2311	1.0233	8.0436	0.0000
Departure (Systematic)	0.1576	1.3958	0.1129	0.9101

For Peer Review Only

FIGURES

Figure 1: Two simple patterns with ϕ=0.50

Figure 2: Flowchart of the proposed PGS

Figure 3: Accuracies achieved from AR process (%)

Figure 4: Accuracies achieved from MA process (%)

Figure 5: Accuracies achieved from ARMA process (%)

Figure 6: Thickness of metallic film in the early stages of the development of an electronic device (Box et al. 2009)

For Peer Review Only

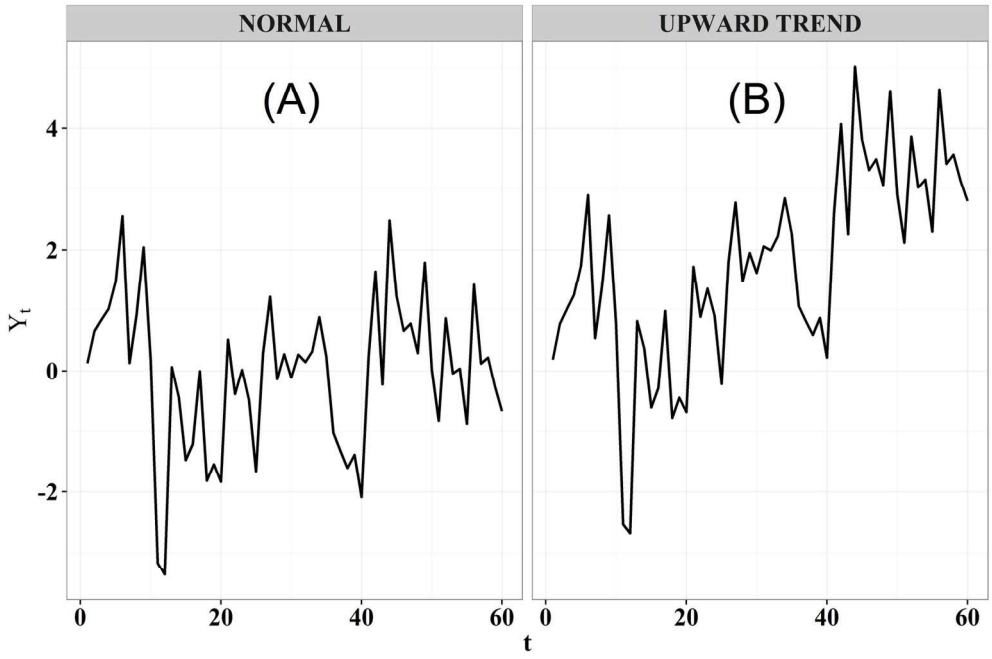

Figure 1: Two simple patterns with ϕ=0.50

152x101mm (300 x 300 DPI)

For Review Only

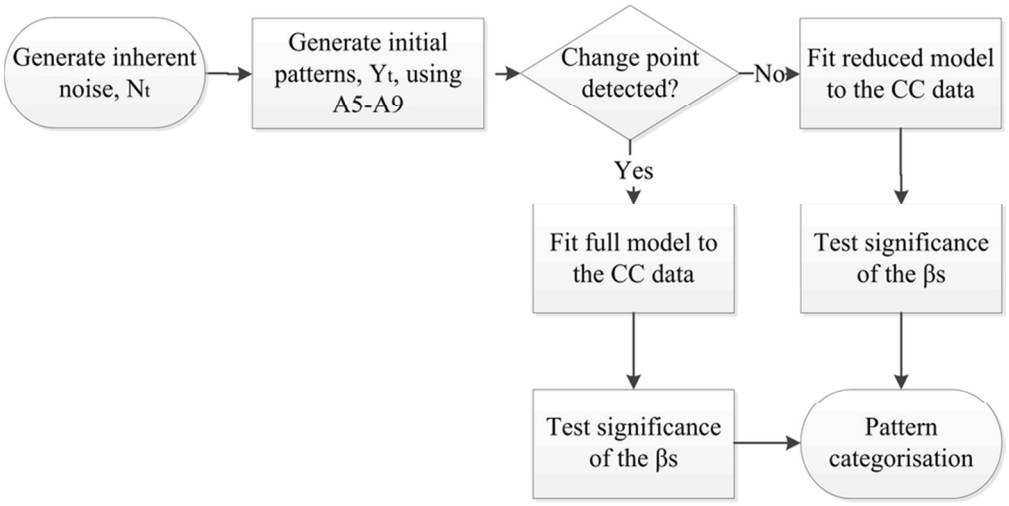

Figure 2: Flowchart of the proposed PGS

81x40mm (300 x 300 DPI)

For Review Only

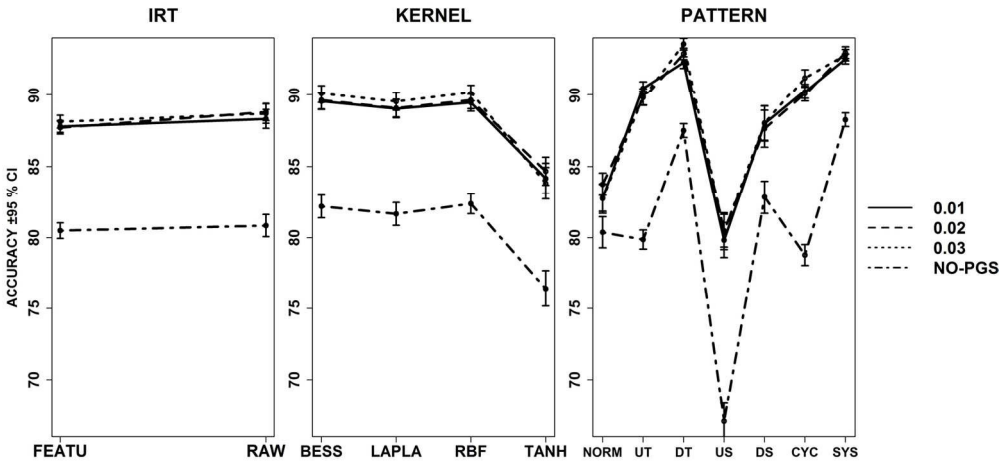

Figure 3: Accuracies achieved from AR process (%)

152x76mm (300 x 300 DPI)

For Review Only

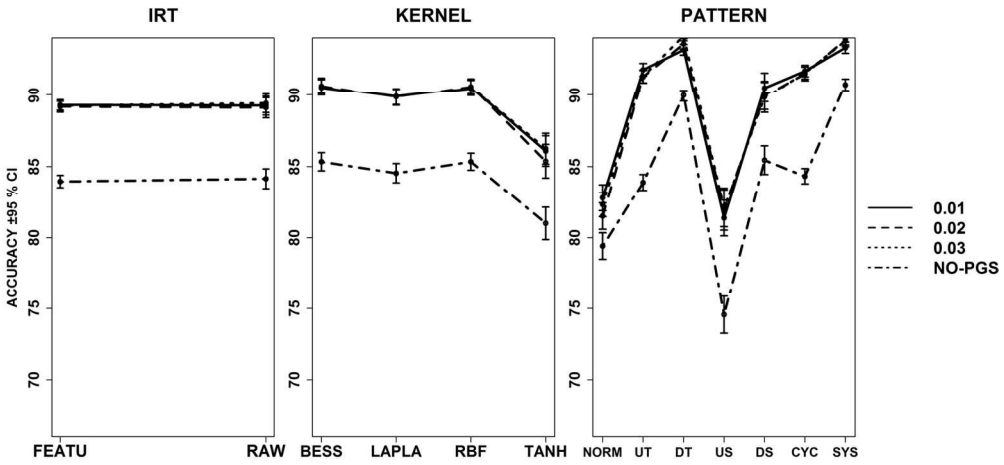

Figure 4: Accuracies achieved from MA process (%)

152x76mm (300 x 300 DPI)

For Review Only

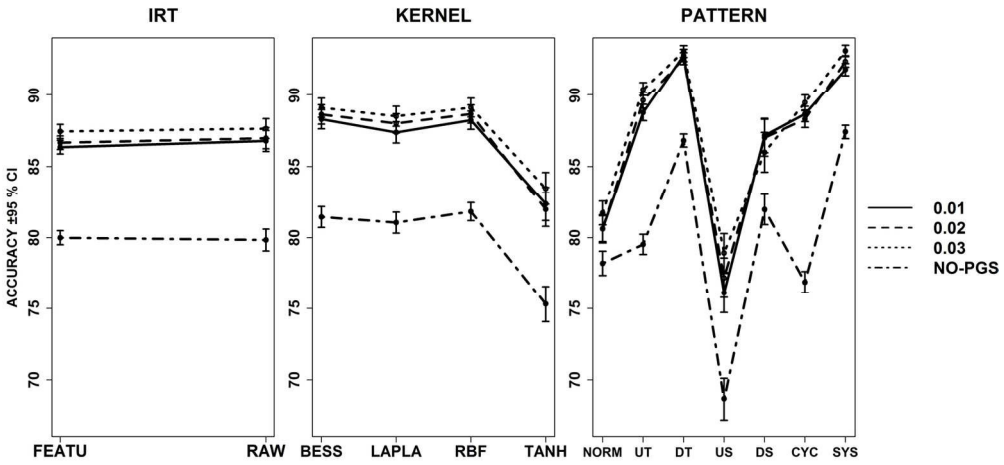

Figure 5: Accuracies achieved from ARMA process (%)

152x76mm (300 x 300 DPI)

For Review Only

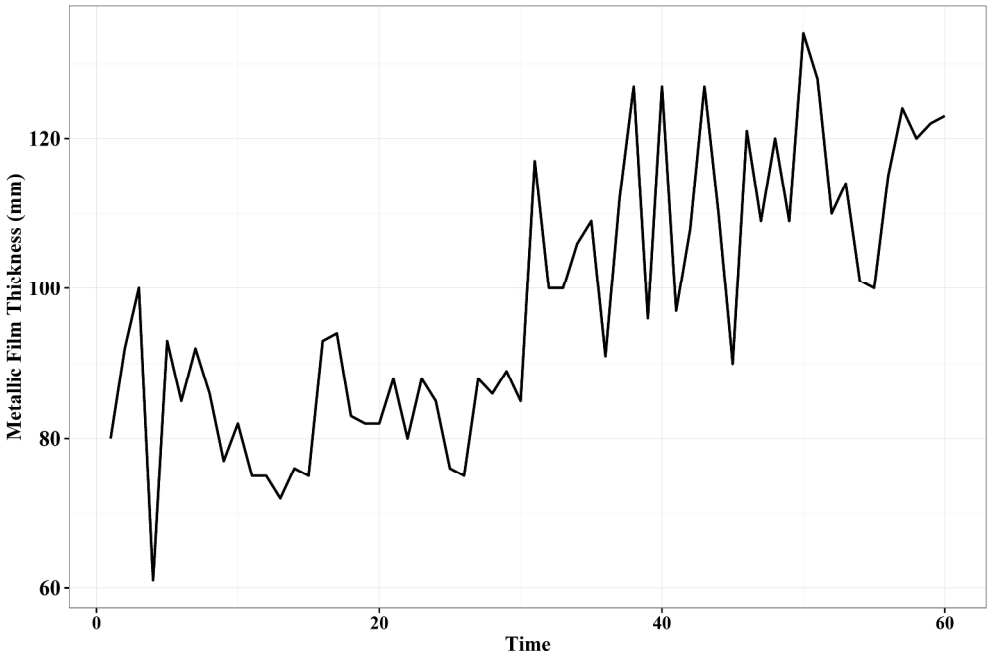

Figure 6: Thickness of metallic film in the early stages of the development of an electronic device (Box et al. 2009)

152x101mm (600 x 600 DPI)

View Only